

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/812,502A

DATE: 05/02/2002

TIME: 10:43:33

Input Set : N:\Crf3\RULE60\09812502A.RAW

Output Set: N:\CRF3\05022002\I812502A.raw

1 <110> APPLICANT: Anderson, Marilyn A.
2 Atkinson, Angela H.
3 Heath, Robyn L.
4 Clarke, Adrienne E.
5 <120> TITLE OF INVENTION: PROTEINASE INHIBITOR, PRECURSOR THEREOF AND GENETIC
6 SEQUENCES ENCODING SAME
7 <130> FILE REFERENCE: 9748B
8 <140> CURRENT APPLICATION NUMBER: 09/812,502A
9 <141> CURRENT FILING DATE: 2001-03-20
11 <150> PRIOR APPLICATION NUMBER: US/09/431,500A
12 <151> PRIOR FILING DATE: 1999-11-01
15 <150> PRIOR APPLICATION NUMBER: 08/454,295
16 <151> PRIOR FILING DATE: 1995-09-01
17 <160> NUMBER OF SEQ ID NOS: 16
18 <170> SOFTWARE: PatentIn Ver. 2.1
20 <210> SEQ ID NO: 1
21 <211> LENGTH: 1104
22 <212> TYPE: DNA
23 <213> ORGANISM: Nicotiana alata
24 <400> SEQUENCE: 1

25	aaggcttgta ccttaaaactg tgatccaaga attgcctatg gagtttgccc gcgttcagaa	60
26	gaaaagaaga atgacgcat atgcaccaac tgttgccgag gcacgaaggg ttgtaagtac	120
27	ttcagtgatg atggaacttt tgtttgtgaa ggagagtctg atcctagaaa tccaaaggct	180
28	tgtaccttaa actgtgatcc aagaattgcc tatggagttt gcccgcgttc agaagaaaag	240
29	aagaatgacg ggatatgcac caactgttgc gcaggcacga aggggtgtaa gtacttcagt	300
30	gatgatggaa cttttgtttg tgaaggagag tctgatccta gaaatccaaa ggcttgcct	360
31	cggaattgcg atccaagaat tgcctatggg atttgccac ttgcagaaga aaagaagaat	420
32	gacggatat gcaccaactg ttgcgcaggc aaaaagggtt gtaagtactt tagtgatgat	480
33	ggaacttttg tttgtgaagg agagtctgat cctaaaaatc caaaggcctg tcctcggaat	540
34	tgtgatggaa gaattgccta tgggatttgc ccactttcag aagaaaagaa gaatgatcgg	600
35	atatgcacca actgctgcgc aggcaaaaag ggttgtaagt actttagtga tgatggaact	660
36	tttgtttgtg aaggagagtc tgatcctaaa aatccaaagg cttgtcctcg gaattgtgat	720
37	ggaagaattg cctatgggat ttgccactt tcagaagaaa agaagaatga tcggatatgc	780
38	acaaactgtt gcgcaggcaa aaagggtgtt aagtacttta gtgatgatgg aacttttgtt	840
39	ggtgaaggag agtctgatcc tagaaatcca aaggcctgtc ctcggaattg tgatggaaga	900
40	attgcctatg gaatttgcgc actttcagaa gaaaagaaga atgacgcat atgcaccaat	960
41	ggttgccgag gcaagaaggg ctgtaagtac tttagtgtat atggaacttt tatttgtgaa	1020
42	ggagaatctg aatatgccag caaagtggat gaatatgttg gtgaagtgga gaatgatctc	1080
43	cagaagtcta aggttgcgtg ttcc	1104
45 <210>	SEQ ID NO: 2	
46 <211>	LENGTH: 1360	
47 <212>	TYPE: DNA	
48 <213>	ORGANISM: Nicotiana alata	

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/812,502A

DATE: 05/02/2002

TIME: 10:43:33

Input Set : N:\Crf3\RULE60\09812502A.RAW

Output Set: N:\CRF3\05022002\I812502A.raw

```

49 <220> FEATURE:
50 <221> NAME/KEY: CDS
51 <222> LOCATION: (97)..(1200)
52 <400> SEQUENCE: 2
53   cgagtaagta tggctgttca cagagttagt ttccttgctc tcctcctctt atttggaatg      60
54   tctctgcttg taagcaatgt ggaacatgca gatgcc aag gct tgt acc tta aac      114
55                                     Lys Ala Cys Thr Leu Asn
56                                     1           5
57   tgt gat cca aga att gcc tat gga gtt tgc ccg cgt tca gaa gaa aag      162
58   Cys Asp Pro Arg Ile Ala Tyr Gly Val Cys Pro Arg Ser Glu Glu Lys
59               10           15           20
60   aag aat gat cgg ata tgc acc aac tgt tgc gca ggc acg aag ggt tgt      210
61   Lys Asn Asp Arg Ile Cys Thr Asn Cys Cys Ala Gly Thr Lys Gly Cys
62               25           30           35
63   aag tac ttc agt gat gat gga act ttt gtt tgt gaa gga gag tct gat      258
64   Lys Tyr Phe Ser Asp Asp Gly Thr Phe Val Cys Glu Gly Glu Ser Asp
65               40           45           50
66   cct aga aat cca aag gct tgt acc tta aac tgt gat cca aga att gcc      306
67   Pro Arg Asn Pro Lys Ala Cys Thr Leu Asn Cys Asp Pro Arg Ile Ala
68               55           60           65           70
69   tat gga gtt tgc ccg cgt tca gaa gaa aag aag aat gat cgg ata tgc      354
70   Tyr Gly Val Cys Pro Arg Ser Glu Glu Lys Lys Asn Asp Arg Ile Cys
71               75           80           85
72   acc aac tgt tgc gca ggc acg aag ggt tgt aag tac ttc agt gat gat      402
73   Thr Asn Cys Cys Ala Gly Thr Lys Gly Cys Lys Tyr Phe Ser Asp Asp
74               90           95           100
75   gga act ttt gtt tgt gaa gga gag tct gat cct aga aat cca aag gct      450
76   Gly Thr Phe Val Cys Glu Gly Glu Ser Asp Pro Arg Asn Pro Lys Ala
77               105           110           115
78   tgt cct cgg aat tgc gat cca aga att gcc tat ggg att tgc cca ctt      498
79   Cys Pro Arg Asn Cys Asp Pro Arg Ile Ala Tyr Gly Ile Cys Pro Leu
80               120           125           130
81   gca gaa gaa aag aag aat gat cgg ata tgc acc aac tgt tgc gca ggc      546
82   Ala Glu Glu Lys Lys Asn Asp Arg Ile Cys Thr Asn Cys Cys Ala Gly
83               135           140           145           150
84   aaa aag ggt tgt aag tac ttt agt gat gat gga act ttt gtt tgt gaa      594
85   Lys Lys Gly Cys Lys Tyr Phe Ser Asp Asp Gly Thr Phe Val Cys Glu
86               155           160           165
87   gga gag tct gat cct aaa aat cca aag gcc tgt cct cgg aat tgt gat      642
88   Gly Glu Ser Asp Pro Lys Asn Pro Lys Ala Cys Pro Arg Asn Cys Asp
89               170           175           180
90   gga aga att gcc tat ggg att tgc cca ctt tca gaa gaa aag aag aat      690
91   Gly Arg Ile Ala Tyr Gly Ile Cys Pro Leu Ser Glu Glu Lys Lys Asn
92               185           190           195
93   gat cgg ata tgc acc aac tgc tgc gca ggc aaa aag ggt tgt aag tac      738
94   Asp Arg Ile Cys Thr Asn Cys Cys Ala Gly Lys Lys Gly Cys Lys Tyr
95               200           205           210
96   ttt agt gat gat gga act ttt gtt tgt gaa gga gag tct gat cct aaa      786
97   Phe Ser Asp Asp Gly Thr Phe Val Cys Glu Gly Glu Ser Asp Pro Lys

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/812,502A

DATE: 05/02/2002

TIME: 10:43:34

Input Set : N:\Crf3\RULE60\09812502A.RAW

Output Set: N:\CRF3\05022002\I812502A.raw

```

98      215      220      225      230
99      aat cca aag gct tgt cct cgg aat tgt gat gga aga att gcc tat ggg      834
100     Asn Pro Lys Ala Cys Pro Arg Asn Cys Asp Gly Arg Ile Ala Tyr Gly
101           235           240           245
102     att tgc cca ctt tca gaa gaa aag aag aat gat cgg ata tgc aca aac      882
103     Ile Cys Pro Leu Ser Glu Glu Lys Lys Asn Asp Arg Ile Cys Thr Asn
104           250           255           260
105     tgt tgc gca ggc aaa aag ggc tgt aag tac ttt agt gat gat gga act      930
106     Cys Cys Ala Gly Lys Lys Gly Cys Lys Tyr Phe Ser Asp Asp Gly Thr
107           265           270           275
108     ttt gtt tgt gaa gga gag tct gat cct aga aat cca aag gcc tgt cct      978
109     Phe Val Cys Glu Gly Glu Ser Asp Pro Arg Asn Pro Lys Ala Cys Pro
110           280           285           290
111     cgg aat tgt gat gga aga att gcc tat gga att tgc cca ctt tca gaa      1026
112     Arg Asn Cys Asp Gly Arg Ile Ala Tyr Gly Ile Cys Pro Leu Ser Glu
113     295           300           305           310
114     gaa aag aag aat gat cgg ata tgc acc aat tgt tgc gca ggc aag aag      1074
115     Glu Lys Lys Asn Asp Arg Ile Cys Thr Asn Cys Cys Ala Gly Lys Lys
116           315           320           325
117     ggc tgt aag tac ttt agt gat gat gga act ttt att tgt gaa gga gaa      1122
118     Gly Cys Lys Tyr Phe Ser Asp Asp Gly Thr Phe Ile Cys Glu Gly Glu
119           330           335           340
120     tct gaa tat gcc agc aaa gtg gat gaa tat gtt ggt gaa gtg gag aat      1170
121     Ser Glu Tyr Ala Ser Lys Val Asp Glu Tyr Val Gly Glu Val Glu Asn
122           345           350           355
123     gat ctc cag aag tct aag gtt gct gtt tcc taagtcctaa ctaataatat      1220
124     Asp Leu Gln Lys Ser Lys Val Ala Val Ser
125     360           365
126     gtagtctatg tatgaaacaa aggcattgcca atatgctctg tcttgccctgt aatctgtaat      1280
127     atggtagtgg agcttttcca ctgcctgttt aataagaaaat ggagcactag tttgttttag      1340
128     ttaaaaaaaaa aaaaaaaaaa      1360
130 <210> SEQ ID NO: 3
131 <211> LENGTH: 368
132 <212> TYPE: PRT
133 <213> ORGANISM: Nicotiana alata
134 <400> SEQUENCE: 3
135     Lys Ala Cys Thr Leu Asn Cys Asp Pro Arg Ile Ala Tyr Gly Val Cys
136     1           5           10           15
137     Pro Arg Ser Glu Glu Lys Lys Asn Asp Arg Ile Cys Thr Asn Cys Cys
138           20           25           30
139     Ala Gly Thr Lys Gly Cys Lys Tyr Phe Ser Asp Asp Gly Thr Phe Val
140           35           40           45
141     Cys Glu Gly Glu Ser Asp Pro Arg Asn Pro Lys Ala Cys Thr Leu Asn
142           50           55           60
143     Cys Asp Pro Arg Ile Ala Tyr Gly Val Cys Pro Arg Ser Glu Glu Lys
144           65           70           75           80
145     Lys Asn Asp Arg Ile Cys Thr Asn Cys Cys Ala Gly Thr Lys Gly Cys
146           85           90           95
147     Lys Tyr Phe Ser Asp Asp Gly Thr Phe Val Cys Glu Gly Glu Ser Asp

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/812,502A

DATE: 05/02/2002

TIME: 10:43:34

Input Set : N:\Crf3\RULE60\09812502A.RAW

Output Set: N:\CRF3\05022002\I812502A.raw

```

148          100          105          110
149    Pro Arg Asn Pro Lys Ala Cys Pro Arg Asn Cys Asp Pro Arg Ile Ala
150          115          120          125
151    Tyr Gly Ile Cys Pro Leu Ala Glu Glu Lys Lys Asn Asp Arg Ile Cys
152          130          135          140
153    Thr Asn Cys Cys Ala Gly Lys Lys Gly Cys Lys Tyr Phe Ser Asp Asp
154          145          150          155          160
155    Gly Thr Phe Val Cys Glu Gly Glu Ser Asp Pro Lys Asn Pro Lys Ala
156          165          170          175
157    Cys Pro Arg Asn Cys Asp Gly Arg Ile Ala Tyr Gly Ile Cys Pro Leu
158          180          185          190
159    Ser Glu Glu Lys Lys Asn Asp Arg Ile Cys Thr Asn Cys Cys Ala Gly
160          195          200          205
161    Lys Lys Gly Cys Lys Tyr Phe Ser Asp Asp Gly Thr Phe Val Cys Glu
162          210          215          220
163    Gly Glu Ser Asp Pro Lys Asn Pro Lys Ala Cys Pro Arg Asn Cys Asp
164          225          230          235          240
165    Gly Arg Ile Ala Tyr Gly Ile Cys Pro Leu Ser Glu Glu Lys Lys Asn
166          245          250          255
167    Asp Arg Ile Cys Thr Asn Cys Cys Ala Gly Lys Lys Gly Cys Lys Tyr
168          260          265          270
169    Phe Ser Asp Asp Gly Thr Phe Val Cys Glu Gly Glu Ser Asp Pro Arg
170          275          280          285
171    Asn Pro Lys Ala Cys Pro Arg Asn Cys Asp Gly Arg Ile Ala Tyr Gly
172          290          295          300
173    Ile Cys Pro Leu Ser Glu Glu Lys Lys Asn Asp Arg Ile Cys Thr Asn
174          305          310          315          320
175    Cys Cys Ala Gly Lys Lys Gly Cys Lys Tyr Phe Ser Asp Asp Gly Thr
176          325          330          335
177    Phe Ile Cys Glu Gly Glu Ser Glu Tyr Ala Ser Lys Val Asp Glu Tyr
178          340          345          350
179    Val Gly Glu Val Glu Asn Asp Leu Gln Lys Ser Lys Val Ala Val Ser
180          355          360          365

182 <210> SEQ ID NO: 4
183 <211> LENGTH: 24
184 <212> TYPE: PRT
185 <213> ORGANISM: Nicotiana alata
186 <400> SEQUENCE: 4
187    Lys Ala Cys Thr Leu Asn Cys Asp Pro Arg Ile Ala Tyr Gly Val Cys
188          1          5          10          15
189    Pro Arg Ser Glu Glu Lys Lys Asn
190          20

192 <210> SEQ ID NO: 5
193 <211> LENGTH: 58
194 <212> TYPE: PRT
195 <213> ORGANISM: Nicotiana alata
196 <400> SEQUENCE: 5
197    Asp Arg Ile Cys Thr Asn Cys Cys Ala Gly Thr Lys Gly Cys Lys Tyr
198          1          5          10          15

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/812,502A

DATE: 05/02/2002

TIME: 10:43:34

Input Set : N:\Crf3\RULE60\09812502A.RAW

Output Set: N:\CRF3\05022002\I812502A.raw

```

199   Phe Ser Asp Asp Gly Thr Phe Val Cys Glu Gly Glu Ser Asp Pro Arg
200               20                25                30
201   Asn Pro Lys Ala Cys Thr Leu Asn Cys Asp Pro Arg Ile Ala Tyr Gly
202               35                40                45
203   Val Cys Pro Arg Ser Glu Glu Lys Lys Asn
204               50                55
206 <210> SEQ ID NO: 6
207 <211> LENGTH: 58
208 <212> TYPE: PRT
209 <213> ORGANISM: Nicotiana alata
210 <400> SEQUENCE: 6
211   Asp Arg Ile Cys Thr Asn Cys Cys Ala Gly Thr Lys Gly Cys Lys Tyr
212       1             5             10             15
213   Phe Ser Asp Asp Gly Thr Phe Val Cys Glu Gly Glu Ser Asp Pro Arg
214               20                25                30
215   Asn Pro Lys Ala Cys Pro Arg Asn Cys Asp Pro Arg Ile Ala Tyr Gly
216               35                40                45
217   Ile Cys Pro Leu Ala Glu Glu Lys Lys Asn
218       50                55
220 <210> SEQ ID NO: 7
221 <211> LENGTH: 58
222 <212> TYPE: PRT
223 <213> ORGANISM: Nicotiana alata
224 <400> SEQUENCE: 7
225   Asp Arg Ile Cys Thr Asn Cys Cys Ala Gly Lys Lys Gly Cys Lys Tyr
226       1             5             10             15
227   Phe Ser Asp Asp Gly Thr Phe Val Cys Glu Gly Glu Ser Asp Pro Lys
228               20                25                30
229   Asn Pro Lys Ala Cys Pro Arg Asn Cys Asp Gly Arg Ile Ala Tyr Gly
230               35                40                45
231   Ile Cys Pro Leu Ser Glu Glu Lys Lys Asn
232       50                55
234 <210> SEQ ID NO: 8
235 <211> LENGTH: 58
236 <212> TYPE: PRT
237 <213> ORGANISM: Nicotiana alata
238 <400> SEQUENCE: 8
239   Asp Arg Ile Cys Thr Asn Cys Cys Ala Gly Lys Lys Gly Cys Lys Tyr
240       1             5             10             15
241   Phe Ser Asp Asp Gly Thr Phe Val Cys Glu Gly Glu Ser Asp Pro Lys
242               20                25                30
243   Asn Pro Lys Ala Cys Pro Arg Asn Cys Asp Gly Arg Ile Ala Tyr Gly
244               35                40                45
245   Ile Cys Pro Leu Ser Glu Glu Lys Lys Asn
246       50                55
248 <210> SEQ ID NO: 9
249 <211> LENGTH: 58
250 <212> TYPE: PRT
251 <213> ORGANISM: Nicotiana alata

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/812,502A

DATE: 05/02/2002
TIME: 10:43:35

Input Set : N:\Crf3\RULE60\09812502A.RAW
Output Set: N:\CRF3\05022002\I812502A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:15; Xaa Pos. 1,2

Seq#:16; Xaa Pos. 1,4,5,21

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/812,502A

DATE: 05/02/2002

TIME: 10:43:35

Input Set : N:\Crf3\RULE60\09812502A.RAW

Output Set: N:\CRF3\05022002\I812502A.raw

L:315 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:318 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:15
L:319 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15 after pos.:0
L:329 M:281 W: Numeric Fields not Ordered, <221> Sort in ascending order!
L:332 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:16
L:335 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:16
L:338 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:16
L:341 M:258 W: Mandatory Feature missing, <220> not found for SEQ ID#:16
L:342 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:0
L:344 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:16